

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 4.Apr.05		3. REPORT TYPE AND DATES COVERED MAJOR REPORT
4. TITLE AND SUBTITLE HYBRIDIZATION OF HYPERSPECTRAL IMAGING TARGET DETECTION ALGORITHM CHAINS			5. FUNDING NUMBERS	
6. AUTHOR(S) CAPT GRIMM DAVID C				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) ROCHESTER INSTITUTE OF TECHNOLOGY			8. PERFORMING ORGANIZATION REPORT NUMBER CI04-1027	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) THE DEPARTMENT OF THE AIR FORCE AFIT/CIA, BLDG 125 2950 P STREET WPAFB OH 45433			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION AVAILABILITY STATEMENT Unlimited distribution In Accordance With AFI 35-205/AFIT Sup 1			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words)				
14. SUBJECT TERMS			15. NUMBER OF PAGES 10	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT	

Hybridization of hyperspectral imaging target detection algorithm chains

David C. Grimm^{a b} and Dr. David W. Messinger^a and Dr. John P. Kerekes^a and Dr. John R. Schott^a

^a Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology Rochester, NY 14623, USA

^b Air Force Institute of Technology, USAF

ABSTRACT

Detection of a known target in an image has several different approaches. The complexity and number of steps involved in the target detection process makes a comparison of the different possible algorithm chains desirable. Of the different steps involved, some have a more significant impact than others on the final result - the ability to find a target in an image. These more important steps often include atmospheric compensation, noise and dimensionality reduction, background characterization, and detection (matched filtering for this research). A brief overview of the algorithms to be compared for each step will be presented.

This research seeks to identify the most effective set of algorithms for a particular image or target type. Several different algorithms for each step will be presented, to include ELM, FLAASH, MNF, PPI, MAXD, the structured background matched filters OSP, and ASD. The chains generated by these algorithms will be compared using the Forest Radiance I HYDICE data set. Finally, receiver operating characteristic (ROC) curves will be calculated for each algorithm chain and, as an end result, a comparison of the various algorithm chains will be presented.

Keywords: ELM, FLAASH, MNF, PPI, MAXD, OSP, ASD, target detection

1. INTRODUCTION

Detection of known substances, or targets, is a very common problem in hyperspectral imaging (HSI). Given a target with a known spectral signature, an algorithm that can decide if and where that target is present in an image is needed. In order to accomplish this, matched filters or target detectors have been developed to accentuate pixels in an image that contain the target.

There are several unique steps to any target detection algorithm. Each of these steps plays a role in determining the overall performance of the target detector. Figure 1 is a flow chart detailing each of these steps for a generic target detection algorithm chain. For each individual step along the way, there are several algorithms that will return an acceptable solution. It is desired to determine the *best* algorithm to accomplish each step. This is where hybridization comes in. A particular combination of algorithms may work well for one image or target but may also perform terribly for a different image or target. Finding an optimal combination of algorithms - or recipes - becomes an intriguing dilemma. In order to best accomplish this, commonly used algorithms for each step must be looked at and compared to each other for many different image/target combinations. This paper will discuss in detail some of the different algorithms used for each step of the target detection process. Background and theoretical explanation of the algorithms will be discussed followed by results, in the form of ROC curves, that will compare several unique algorithm chains.

1.1. Data Set

The data set being used for this paper is the "Forest Radiance 1" scene collected by the HYDICE sensor. The image dimensions are 320x1280 with 145 spectral bands covering the spectrum from approximately 350nm to 2500nm. Two targets will be used to compare the algorithm chains, one of which is "easy" to find and one of which is slightly more difficult to find. There are many man made targets placed into this scene. In order to determine the overall difficulty of detection, the Spectral Angle Mapper (SAM) detection algorithm was run for each target in the scene. SAM is a very basic target detection algorithm that relies on the spectral angle

20050412 024

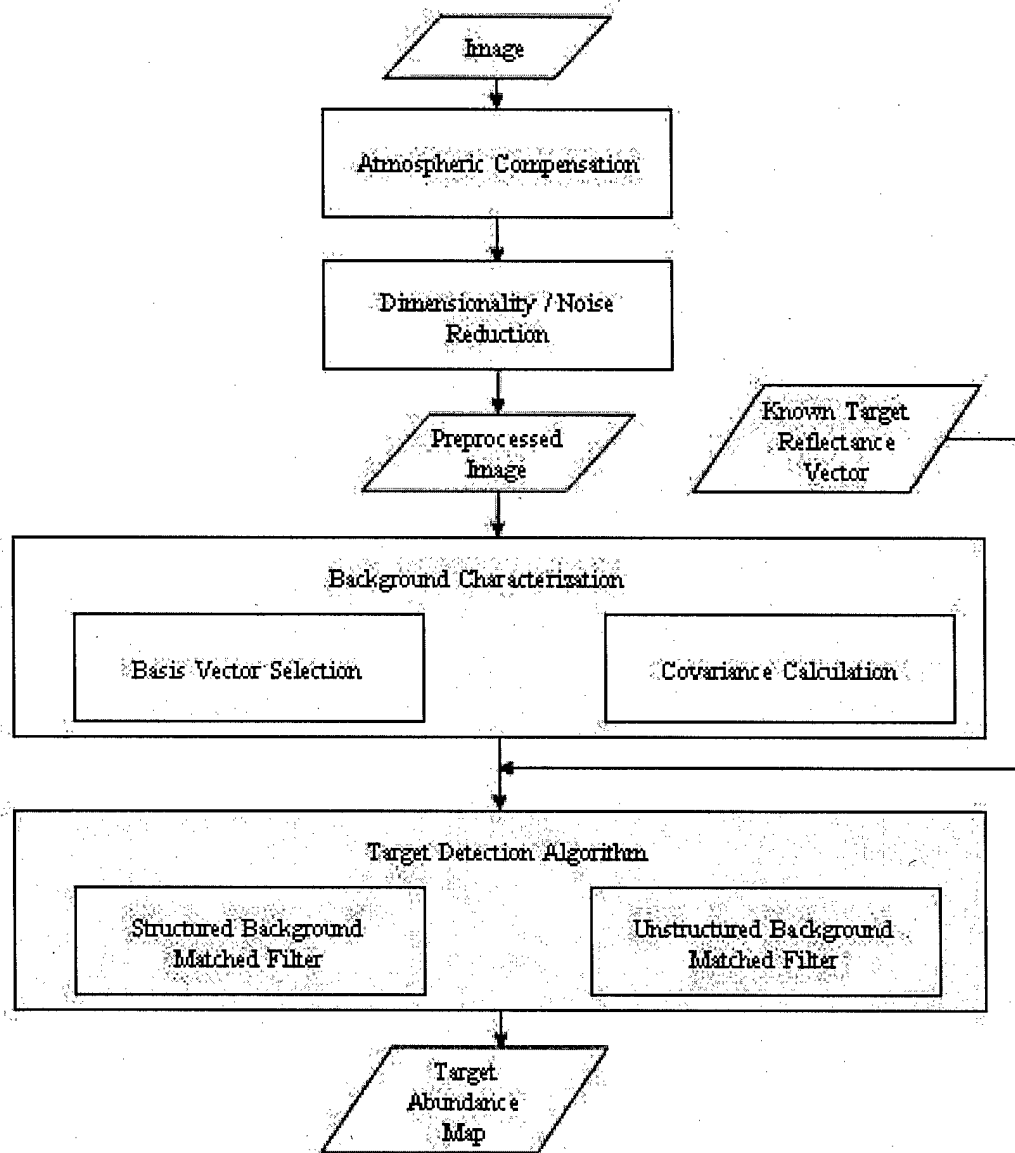


Figure 1. Typical target detection algorithm flow chart

between the target and pixel in question to determine if the target is present. Equation 1 is the mathematical representation of this detector

$$T_{SAM}(x) = \frac{d^T x}{(d^T d)^{1/2} (x^T x)^{1/2}} \quad (1)$$

where d is the target spectra and x is the pixel in question. For the purposes of this paper, the SAM results were only used to determine the targets to be used for comparison and to provide a baseline ROC curve (i.e. if an algorithm chain cannot outperform SAM it is not really worth looking at).

A second basic matched filter is the Generalized Likelihood Ratio Test (GLRT). The results returned by this matched filter will also be used for comparison purposes. The GLRT relies on an unstructured, or stochastic, representation of the background in the image in the form of a covariance matrix. The equation for GLRT is shown in Equation 2:

$$T_{GLRT}(x) = \frac{(d^T \Sigma^{-1} x)^2}{(d^T \Sigma^{-1} d)(1 + x^T \Sigma^{-1} x)} \quad (2)$$

where d and x are the same as Equation 1 and Σ is the covariance matrix used. The GLRT results shown were generated using the covariance of the “tree” class of pixels in the image.

Based on the results of these two matched filters, the “easy” target, Target A, had an average false alarm rate (AFAR) of 0.003313271 for SAM and 0.009138755 for the GLRT. The lower the AFAR, the easier the target is to detect. The AFAR is defined as the average number of false alarms encountered for each target pixel detected. The “difficult” target, Target B had AFARs of 0.023396362 and 0.076723623 for SAM and GLRT respectively. Figure 2 is a simple three band RGB image of the data set and also shows the spatial locations of the targets being used.

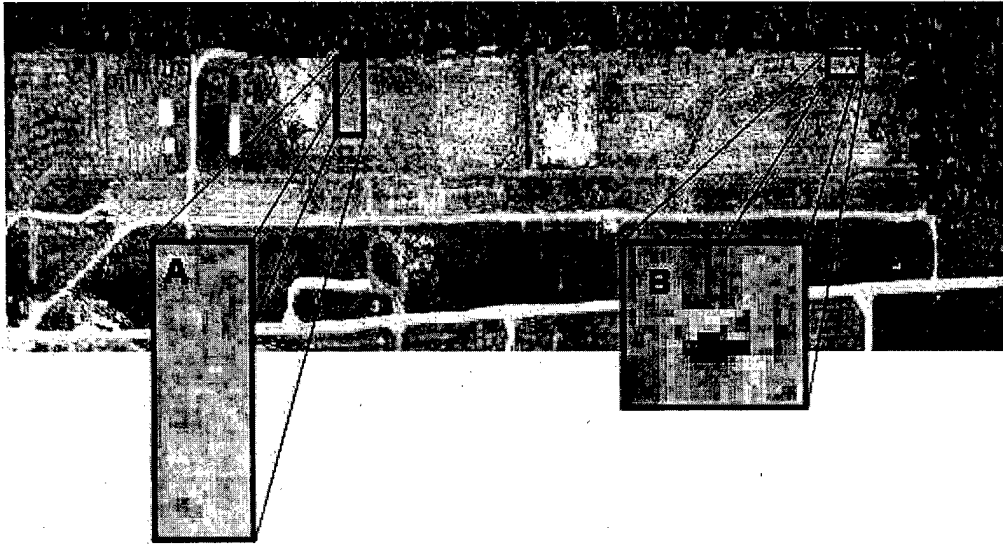


Figure 2. Spatial locations of “easy” target A and “difficult” target B in the Forest Radiance 1 scene

By simply looking at the image, it can be seen that, barring the planted targets, the scene is very uncluttered with very little man-made material. There are only about six or seven different classes of pixels present in the image and several of those are rather similar (i.e. light and dark grass are considered unique classes). Secondly, there are no “concealed” target pixels, all of them are out in the open and completely uncovered. There are both fully resolved and sub-pixel target pixels for both targets. For the purposes of this paper, all planted

targets in the image except for Targets A and B have been masked out for the detection results and background characterization. All pixels with any amount of target present are searched for and counted as part of the probability of detection statistic.

2. ATMOSPHERIC COMPENSATION

Atmospheric compensation is the process of "taking the atmosphere out of the image". The known spectral signature of the targets consist of measured reflectance values. These measurements cannot take into account the effects of looking through thousands of feet worth of atmosphere. The ultimate goal of the atmospheric compensation process is to retrieve reflectance from the radiance recorded by the sensor. The following sections will briefly outline the theory behind the two atmospheric compensation methods used.

2.1. Empirical Line Method (ELM)

The ELM relies on ground truth inasmuch as at least two different regions in the image (preferably one dark and one light across each wavelength) must have a known reflectivity. These regions must be at least one pixel large. Ideally, the regions should have corresponding ground truth reflectance spectra taken at the same time as the image. If such regions are not present in the image, an educated guess can be made by selection of regions for which an approximate reflectance can be created. For example, a white cloud may have a reflectivity of about 90% across all wavelengths. Obviously, there is a great opportunity for error introduction if estimates have to be made, but that can be limited if they are made smartly. The Forest Radiance 1 image has calibration panels present and the values measured from those panels were used to determine the ELM compensation.

Once the two (or more) regions have been selected, a line is fit through the radiance (or digital counts) vs reflectivity points in each band. ELM assumes a linear relationship between the radiance (or digital counts) and the reflectivity. Mathematically, this relationship is expressed Equation 3

$$L = mR(\lambda) + b \quad (3)$$

where L is the observed radiance (or digital count), m is the slope of the line through the ground truth points, $R(\lambda)$ is the wavelength dependent reflectance, and b is the radiance (or digital count) value that represents zero reflectance. Figure 3 illustrates this relationship graphically.

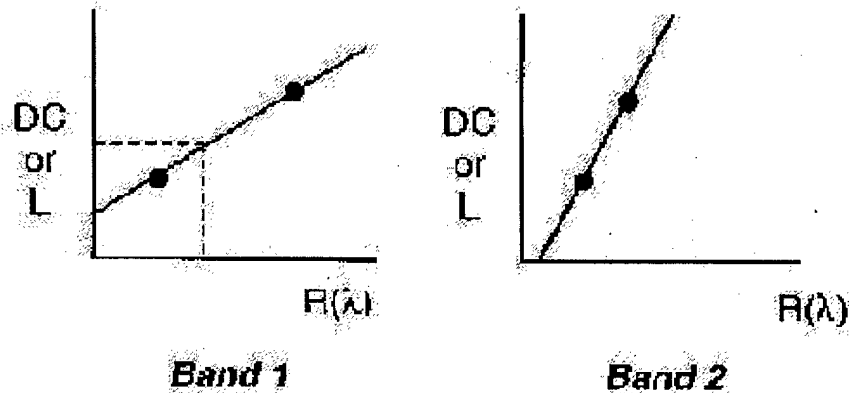


Figure 3. Graphical representation of ELM for two bands. The marked points are the "ground truth" points

Now, the observed radiance in every pixel in the image can be converted into reflectance using the linear regression coefficients (m and b) obtained via Equation 3. It is important to note that ELM can work with both radiance or digital counts without calibration.

2.2. Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes (FLAASH)

FLAASH is a commercially available atmospheric compensation algorithm developed by the Air Force Research Laboratory, Space Vehicles Directorate (AFRL/VS) to support the analysis of visible to shortwave infrared (0.4 μm to 2.4 μm) hyperspectral and multispectral sensors [Berk et al., 2002]. Using atmospheric inversion, FLAASH is able to determine the aerosol and water vapor content of the atmosphere in an image as well as the surface pressure.

Equation 4 [Berk et al., 2002] is the sensor reaching radiance (L^*) equation utilized by FLAASH:

$$L^* = \frac{A\rho}{1 - \rho_e S} + \frac{B\rho_e}{1 - \rho_e S} + L_a^* \quad (4)$$

where ρ is the pixel surface reflectance, ρ_e is an average surface reflectance for the area surrounding the pixel in question, S is the spherical albedo of the atmosphere (accounting for the skylight photons), L_a^* is the up-welled radiance, A and B are surface independent coefficients that vary based on the atmospheric and geometric conditions. All of the variables are implicitly wavelength dependent.

The per-pixel A , B , S , and L_a^* variables are all determined empirically via MODTRAN. Once these values have been extracted, Equation 5 is used to solve for ρ_e [Berk et al., 2002]:

$$L_e^* = \frac{(A + B)\rho_e}{1 - \rho_e S} + L_a^* \quad (5)$$

where L_e^* is a spatially averaged sensor reaching radiance value. Once ρ_e has been determined, Equation 4 can be solved for reflectance (ρ).

3. MAXIMUM NOISE FRACTION (MNF)

Only one algorithm for dimensionality/noise reduction will be looked at in this paper, the Maximum Noise Fraction (MNF) [Green et al., 1988]. Essentially, the MNF algorithm is the same as the Principle Components Analysis process with some “pre-processing” steps added. For the MNF transform, the covariance matrix of the noise present is needed. In order to obtain this (since it is not usually known), a spectrally flat field can be used. Once a flat field is identified within an image, the mean can be subtracted from the observed values of that field. The result is a representation of the noise present in the image. Once this “noise image” is found, its covariance matrix can be calculated. If a flat noise field cannot be found, the ENVI implementation of MNF allows for the noise statistics to be computed from the image.

The next step in this process is basically to perform PCA on the noise image to de-correlate the noise. This is done by calculating the eigenvectors of the noise image. By definition, the resulting eigenvectors are orthogonal to each other. The original image can now be “passed through” this orthogonal transformation. After this is done, the orthogonally transformed image can be whitened, or normalized, by the eigenvalues of the noise covariance matrix. The final result of all these steps is an image in which all of the noise present is orthogonal and identical. Now, the PCA previously described can be performed.

The advantage of the MNF transform is that the “pre-processing” steps force the assumption used in PCA of variance representing information to be true. There is now equal variance present due to noise in the image. The bands are rank ordered by the signal to noise ratios and the highest PC bands are retained. The result is a greatly reduced dimension space with little noise to interfere with target detection.

4. ENDMEMBER SELECTION

There are two overarching ways to characterize the background of an image: structured or unstructured. While an unstructured matched filter (GLRT) was used in determining the targets used and for comparison purposes, this paper will focus on matched filters that employ a structured background. Two related endmember selection algorithms will be compared in this paper.

4.1. Pixel Purity Index (PPI)

The first end member selection algorithm is the Pixel Purity Index (PPI) method [Boardman et al., 1995]. Using PPI, all of the pixel values in an image are projected onto randomly selected vectors. Each time a projection is accomplished, the extrema points are noted. Spectrally pure pixels will consistently be extrema points of these random projections. A major assumption of this method is that spectrally pure pixels in the image are end members. This is generally not a bad assumption. This method is also rather susceptible to noise, so steps (*e.g.* thresholding) should be taken to help lessen the negative effects of noisy pixels. Figure 4 is an example projection of data points. ENVI uses a threshold factor to help mitigate noise interference. The threshold factor determines the amount of pixels other than strictly the extrema pixels are considered - pixels within the threshold factor of the extrema pixels are considered extrema as well.

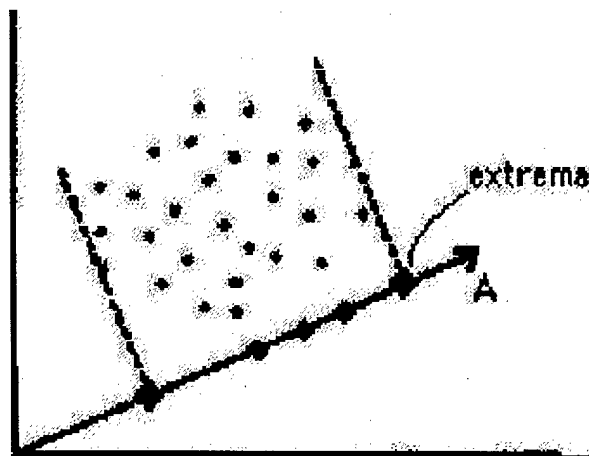


Figure 4. Example PPI projection [Boardman et al., 1995]

For this paper, the PPI implementation in ENVI was used. Using ENVI, the PPI endmember selection process is not fully automated and does require a “man-in-the-loop”. Once the desired number of iterations (normally several thousand) are completed, every pixel that was found as an extrema at least once is returned. It is up to the user to determine how many times a pixel must be found to be considered significant. Once this is completed, the significant pixels must be clustered. The amount of spectral clusters determine the number of endmembers in the image. All of this is done by the user. Finally, an endmember is extracted by taking the mean of each cluster.

4.2. MAXD

The second end member selection technique is MAXD [Lee, 2003]. Similar to PPI, MAXD also uses projection to help sort out the end members of an image. Given a set of data points, MAXD begins by selecting the points that have the maximum and minimum Euclidean distance from the origin. These two points are the first two end members. Once these points are selected, all of the data points are projected onto a vector orthogonal to both of the vectors already selected. This process is repeated until the desired number of end members is found. Figure 5 pictorially explains how this process works.

Unlike PPI, MAXD is used as a fully automated process where the only user input is the number of endmembers to be found. The clustering step in PPI is not needed for MAXD.

5. MATCHED FILTERS

The ultimate goal of the target detection algorithm chain is *target detection*. The means to accomplish target detection in this case are matched filters. Two similar but different structured background matched filters compared in this paper are the Orthogonal Subspace Projector (OSP) and the Adaptive Subspace Detector (ASD).

5.1. Orthogonal Subspace Projector (OSP)

The first matched filter discussed OSP. Equation 6 represents the OSP detector.

$$T_{OSP}(x) = \frac{dP_b x}{d^T P_b d} \quad (6)$$

where d is the desired target vector and all of the other variables are the same as those defined for Equation 7.

OSP projects each pixel into a space orthogonal to the background of the image which is represented by the endmembers extracted. This, in theory, forces the target pixels to “stick out” from the background. The pixels that most closely match the given target spectra are given a high T_{OSP} value.

5.2. Adaptive Subspace Detector (ASD)

A second implementation of a structured background model detector the ASD. This detector is represented mathematically by Equation 7:

$$T_{ASD}(x) = \frac{x^T (P_b - P_S)x}{x^T P_S x} \quad (7)$$

where $P_b = I - U(U^T U)^{-1}U^T$ and $P_S = I - S(S^T S)^{-1}S^T$ and $S \triangleq [dU]$. In words, $P_b x$ is the projection of the pixel, x , onto the vector orthogonal to the background subspace.

As with the OSP matched filter, P_b is the projection of the pixel onto a space orthogonal to the background. The major difference between OSP and ASD is the P_S term in ASD. This is the projection of the pixel onto a space orthogonal to the background *and* the target. The difference between these two projections is a measure of just how “target-like” the pixel is.

6. RESULTS

6.1. Experiment

The first step in the algorithm chains is atmospheric compensation. The “Forest Radiance 1” was atmospherically compensated using both the ELM and FLAASH methods discussed in Section 2. The second step is noise and dimensionality reduction. This is accomplished using the MNF transform outlined in Section 3. The implementation available in ENVI was used for this paper. The option to “estimate noise statistics from data” provided by ENVI was used to obtain a noise spectrum for the transform. For both atmospherically compensated images, over 98% of the total variability was retained resulting in keeping 136 and 137 bands for the ELM image and FLAASH image respectively.

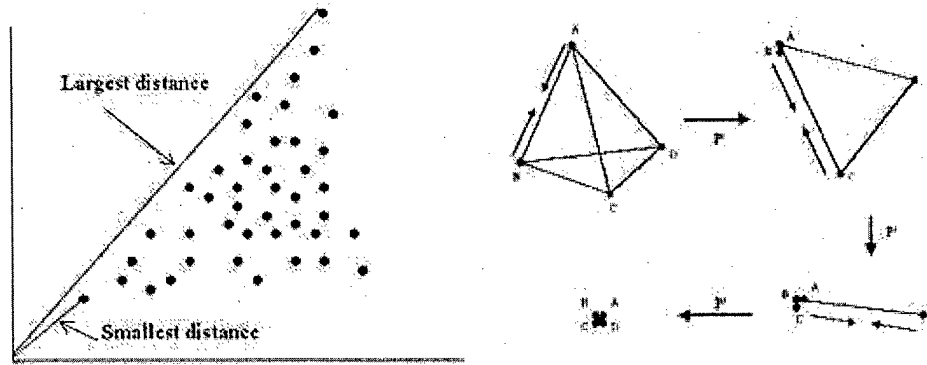


Figure 5. MAXD end member selection algorithm [Lee, 2003]

As mentioned in Section 4, PPI and MAXD were used to extract endmembers to characterize the background of the image. The procedure for usage of the PPI application in ENVI discussed in Section 4.1 was followed using 30,000 iterations and a threshold factor of 2. Seven different classes were found in the clustering step, resulting in a background matrix consisting of 7 endmembers. The MAXD algorithm is not currently available in ENVI, so IDL was used. The only user input necessary for MAXD is the number of endmembers to be found, for this image, 20 endmembers were found.

Finally, each image/background combination was used with the matched filters presented in Section 5. Figure 6 is the ROC curves that resulted from each combination for the two targets as well as the SAM and GLRT results for comparison purposes.

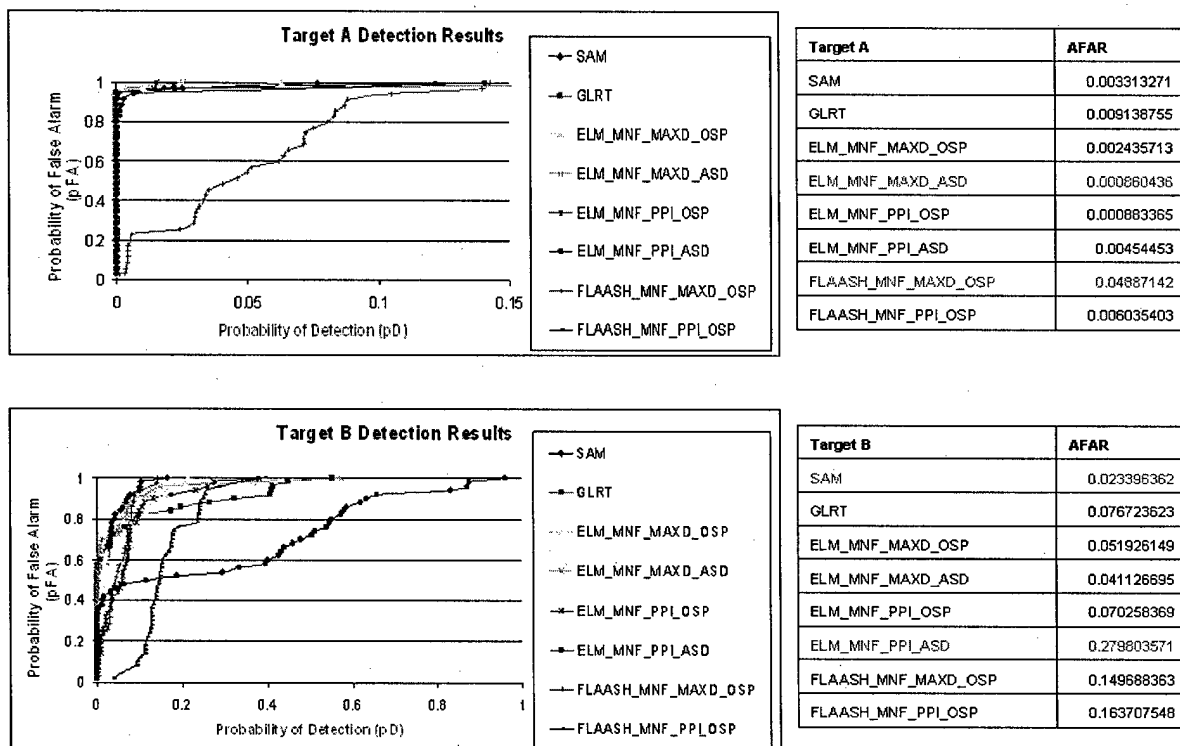


Figure 6. Resulting ROC curves and AFAR for the various algorithm chains

6.2. Conclusions

Overall, there were no results that seemed particularly telling. For Target A, the “easy” target, all of the chains performed comparably with the exception of one, the FLAASH-MNF-MAXD-OSP chain. Looking at the Target B results, this chain underachieved a little there as well. This logically leads to the conclusion that that group of algorithms do not work together well. On the other hand, there did not seem to be one particularly stellar performing chain of algorithms either. Quite surprisingly, the detector that was able to detect the more difficult Target B was SAM. SAM is not dependent on any background characterization and is generally considered the absolute baseline for detection performance.

It was expected that the chains using ELM would have greatly outperformed the chains using FLAASH due to the fact that good ground truth went into the ELM calculation. While this is generally true for the two targets, a much greater difference would have been expected. It would be interesting to repeat the experiment using ELM on the scene without using the collected ground truth and see if the results change drastically. Another intriguing question is what the results would look like for a much more cluttered (i.e. urban) scene.

7. SUMMARY

This paper is intended to be an exploration of different algorithm chains, lending itself to more ongoing research looking at many more unique chains. The goal of the research is to try to *pre*-determine the *best* set of algorithms to solve a target detection problem for any given image and/or target. A few future work problems were brought up in Section 6.2, and many more have arisen based on this small piece of the much larger pie.

References

- A. Berk, S. M. Adler-Golden, A. J. Ratkowski, G. W. Felde, G. P. Anderson, M. L. Hoke, T. Cooley, J. H. Chewynd, J. A. Gardner, M. W. Matthew, L. S. Bernstein, P. K. Acharya, D. Miller, and P. Lewis. Exploiting MODTRAN radiation transport for atmospheric correction: The FLAASH algorithm. In *Proceedings of the Fifth International Conference on Information Fusion*, volume 2, July 2002.
- Joseph W. Boardman, Fred A. Kruse, and Robert O. Green. Mapping target signatures via partial unmixing of AVIRIS data. In *Fifth JPL Airborne Earth Science Workshop*, volume 1 of *JPL Publication 95-1*, pages 23–26, 1995.
- Andrew A. Green, Mark Berman, Paul Switzer, and Maurice D. Craig. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. In *IEEE Transactions on Geoscience and Remote Sensing*, volume 26, pages 65–74, January 1988.
- Kyungsuk Lee. *A Subpixel Target Detection Algorithm for Hyperspectral Imagery*. PhD thesis, Rochester Institute of Technology, May 2003.